

DEL DISEÑO CONVENCIONAL AL DISEÑO OPTIMO. POSIBILIDADES Y VARIANTES Parte II. Optimización multiobjetivo y sensibilidad de la solución óptima

SANTIAGO HERNANDEZ

*Departamento de Ingeniería Mecánica,
Centro Politécnico Superior de Ingenieros,
Universidad de Zaragoza, María de Luna, 3,
50015 Zaragoza, España*

RESUMEN

Esta segunda parte comienza describiendo como pueden tratarse varias funciones objetivo dentro de un mismo proceso de optimización, a fin de acercar el modelo matemático a situaciones más reales del diseño en la ingeniería. En esta optimización multiobjetivo una vez más debe tratarse con el inconveniente de la no convexidad del problema; por ello se comentan métodos que pueden tratar con esta situación y se aporta un ejemplo estructural. El trabajo termina comentando los métodos existentes para abordar la modificación de la solución óptima cuando se cambia alguno de los parámetros fijos de un problema de optimización y se resuelve una situación de este tipo mediante uno de ellos.

SUMMARY

This second part starts describing how to deal with several objective functions in an optimization problem. This means a step forward in order to close up mathematical models of optimization and real life engineering problems. This approach, called multicriteria optimization, presents big inconvenients when nonconvexity arises, and methods to cancel out this situation are analyzed. This work ends explaining what happen when fixed parameters in the optimization problem change and how a new optimal solution for the new parameter values may be obtained without starting again the optimization procedure. The different techniques are presented and a structural problems is solved to show one of these methods performance.

Recibido: Septiembre 1988

METODOS DE OPTIMIZACION MULTIOBJETIVO

Muchas veces la tarea del ingeniero es encontrar un diseño en el que se quiere optimizar más de una función objetivo. A consecuencia de ello aparece un vector \mathbf{F} cuyos componentes son las distintas propiedades a optimizar. Este planteamiento se denomina optimización multiobjetivo^{1,2} y se escribirá:

$$\min \mathbf{F} = |f_1(\mathbf{X}), \dots, f_K(\mathbf{X})| \quad (1.a)$$

sujeto a

$$g_j(\mathbf{X}) \leq 0 \quad j = 1, \dots, m \quad (1.b)$$

La intención del diseñador sería encontrar un vector de variables \mathbf{X} tal que:

$$f_k(\mathbf{X}^*) = \min f_k(\mathbf{X}) \quad k = 1, \dots, K \quad (2)$$

Sin embargo esto no suele producirse y habitualmente las funciones objetivo varían de forma contraria. En la Figura 1 se observa un problema de una variable y dos funciones objetivo. El mínimo de cada uno de ellos se produce en un punto distinto y el objetivo de (2) no es alcanzable.

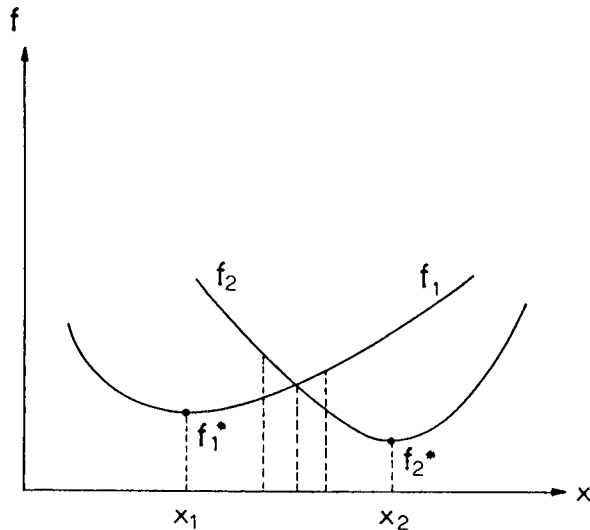


Figura 1. Optimización con una variable y dos funciones objetivo.

En estos casos resulta necesario acudir a los puntos definidos como mínimos de Pareto³ que son aquellos puntos \mathbf{X}^P para los que no existe ningún punto \mathbf{X} tal que:

$$a) f_k(\mathbf{X}) \leq f_k(\mathbf{X}^P) \quad k = 1, \dots, K \quad (3.a)$$

$$b) f_k(\mathbf{X}) < f_k(\mathbf{X}^P) \quad \text{para una función objetivo, al menos.} \quad (3.b)$$

La característica de este tipo de mínimo es que al separarse de él para disminuir el valor de alguna función objetivo, se incrementa, al menos, una de las restantes funciones objetivo. En la Figura 1 todos los valores de la variable que cumplen $x_1 \leq x \leq x_2$ son mínimos de Pareto.

En los problemas con varias variables los mínimos de Pareto son numerosos y se encuentran en el contorno, o en los puntos del dominio donde son tangentes dos o más funciones objetivo.

En la optimización multiobjetivo resulta muy importante representar el problema en el espacio de las funciones objetivo. Ello puede hacerse eliminando las variables de diseño a partir de las condiciones activas. El resultado es una o varias curvas que definen un contorno cerrado del que formarán parte los mínimos de Pareto existentes. En la Figura 2 se presenta un caso con dos variables y dos funciones objetivo. Los mínimos de Pareto aparecen dibujados con línea discontinua.

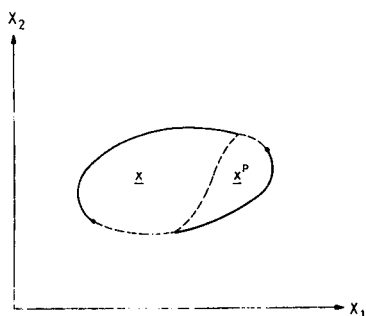


Figura 2a. Región de diseño en el espacio de las variables de diseño.

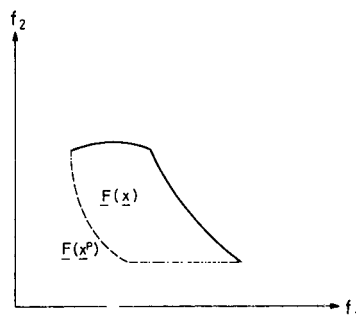


Figura 2b. Región de diseño en el espacio de las funciones objetivo.

La línea horizontal de la Figura 2b está compuesta por *mínimos débiles de Pareto* que son aquellos puntos X^P , para los que no hay ningún punto X que cumpla:

$$f_k(X) < f_k(X^P) \quad k = 1, \dots, K \quad (4)$$

Supóngase que en la estructura de tres barras de la Figura 3 se plantea una optimización considerando dos funciones objetivo, el volumen V y el movimiento vertical u_2 del nudo. Si se incluyen limitaciones en las variables y se escriben únicamente las condiciones activas, el problema podrá escribirse como:

$$\min V = (2\sqrt{2}x_1 + x_2)l \quad (5.a)$$

$$\min u_2 = \frac{Pl}{E(x_1 + \sqrt{2}x_2)} \quad (5.b)$$

sujeto a

$$\sigma_1 = \frac{x_2 + \sqrt{2}x_1}{2x_1x_2 + \sqrt{2}x_1^2}P \leq 2 \quad (5.c)$$

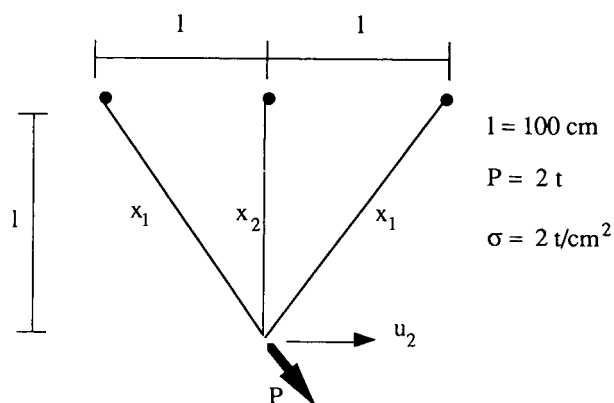


Figura 3. Ejemplo de estructura de tres barras.

$$0 \leq x_1 \leq 1.414 \quad (5.d)$$

$$0 \leq x_2 \leq 2.828 \quad (5.e)$$

El espacio de diseño resultante aparece en la Figura 4, y los mínimos de Pareto se encuentran entre los puntos \mathbf{X}_E , \mathbf{X}_A , \mathbf{X}_B del contorno.

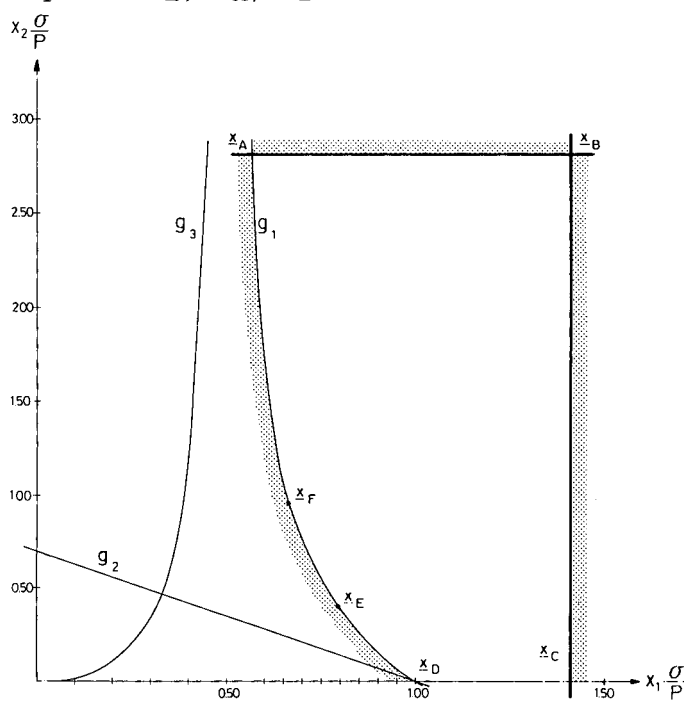


Figura 4. Región de diseño de la optimización multicriterio de la estructura de tres barras.

Al representar el problema en el espacio de las funciones objetivo se obtiene la Figura 5. Comparando ambas figuras se observa que el diseño que hace mínimo a u_2 es \mathbf{X}_B y el de menor volumen es \mathbf{X}_E .

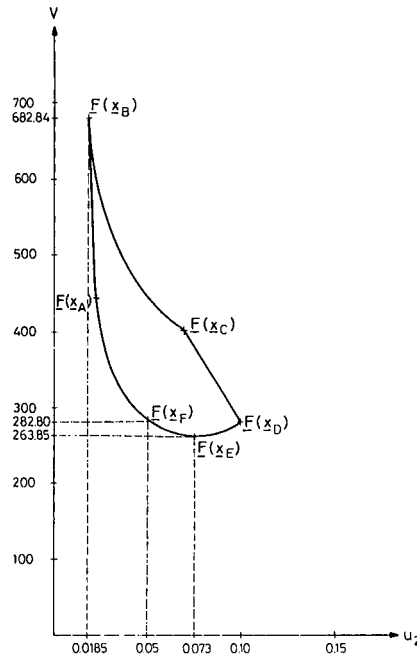


Figura 5. Región de diseño en el espacio de las funciones objetivo de la estructura de tres barras.

Dado que habitualmente no existirá un mínimo común para todas las funciones objetivo no existe una minimización en sentido estricto en este tipo de problemas. Por ellos la tarea de un diseñador será intentar identificar el mayor número de mínimos de Pareto posible y a la vista de ello elegir aquél que le parece que satisface de forma más adecuada el conjunto de funciones objetivo. Por ello el diagrama de flujo del proceso será el mostrado en la Figura 6:

Existen diferentes métodos⁴ para obtener el conjunto de mínimos de Pareto. Uno muy común consiste en definir un vector λ de coeficientes de ponderación y linealizar así la función objetivo.

$$F = \lambda^T F = \sum_{k=1}^K \lambda_k f_k \quad (6)$$

habitualmente los componentes de λ están normalizados

$$\sum_{k=1}^K \lambda_k = 1 \quad (7)$$

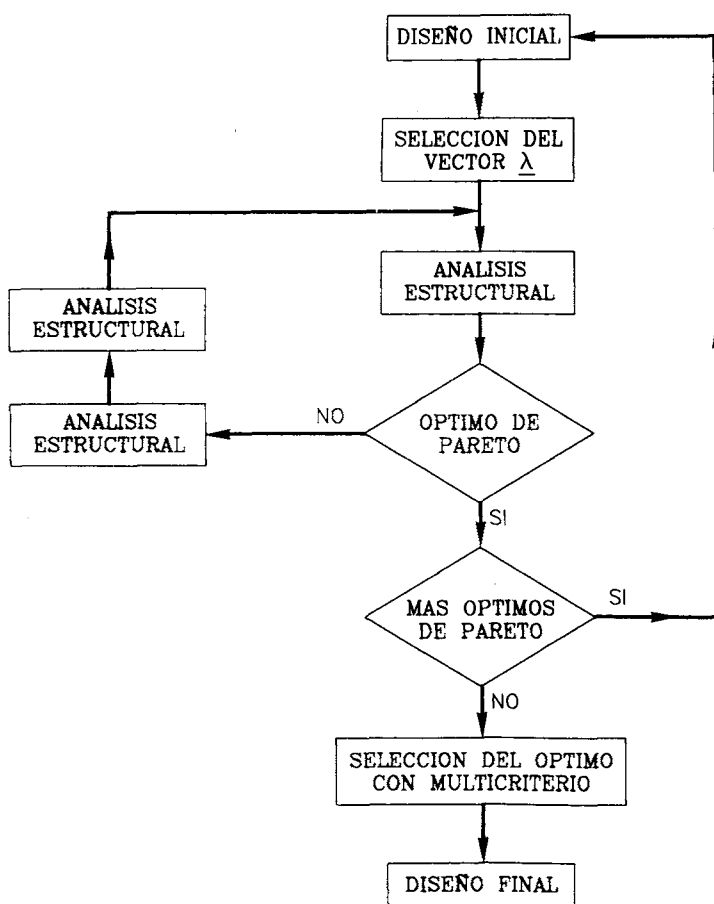


Figura 6. Diagrama de flujo en la optimización multicriterio.

En la Figura 7a) se puede observar que al cambiar los coeficientes de λ se obtienen los puntos comprendidos entre f_1^* y f_2^* . Sin embargo si el contorno de las soluciones no es convexo, como sucede en la Figura 7b) no es posible obtener los puntos situados entre A y B . Otro inconveniente es que si los órdenes de magnitud de las funciones objetivo es muy distinto resulta difícil encontrar mínimos de Pareto suficientemente diferentes, incluso haciendo grandes cambios en los componentes del vector λ .

Para evitar estas desventajas existe una técnica denominada *min-max* en que las funciones objetivo se normalizan en el rango $[0,1]$.

$$\bar{f}_k = \frac{f_k - \min f_k}{\max f_k - \min f_k} \quad k = 1, \dots, K \quad (8)$$

el nuevo vector de funciones objetivo es

$$\bar{\mathbf{F}} = |\bar{f}_1, \dots, \bar{f}_K| \quad (9)$$

A continuación se asigna un coeficiente λ_k a cada función y se procede a resolver el problema.

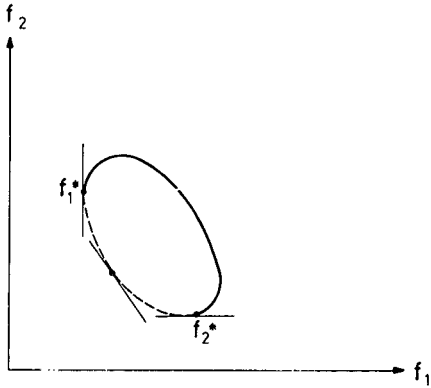


Figura 7a. Región de diseño convexa en el espacio de las funciones objetivo.

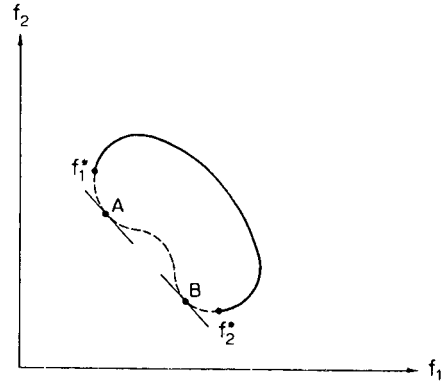


Figura 7b. Región de diseño no convexa en el espacio de las funciones objetivo.

$$\min(\max \lambda_k \bar{f}_k) \quad k = 1, \dots, K \quad (10)$$

Todos los mínimos de Pareto pueden ser obtenidos de esta forma. El problema de (10) se resuelve definiendo una variable γ y estableciendo la siguiente formulación.

$$\min \gamma \quad (11.a)$$

sujeto a

$$\lambda_k \bar{f}_k \leq \gamma \quad k = 1, \dots, K \quad (11.b)$$

$$g_j(\mathbf{X}) \leq 0 \quad j = 1, \dots, m \quad (11.c)$$

La interpretación geométrica del problema puede observarse en la Figura 8 donde para cada vector λ se obtiene un vector que une el origen de coordenadas con el recinto de soluciones.

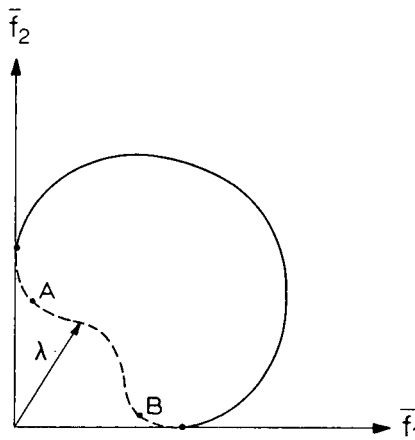


Figura 8. Justificación gráfica del método min-max.

SENSIBILIDAD DE LA SOLUCION OPTIMA

Se recordará que el planteamiento habitual del diseño óptimo con una función objetivo F es de la forma

$$\min (o \max) \quad F(\mathbf{X}, \mathbf{p}) \quad (12)$$

la solución es función del vector de parámetros fijos \mathbf{p}

$$\mathbf{X}^* = \mathbf{X}^*(\mathbf{p}) \quad (13)$$

y en consecuencia el valor de la función objetivo es

$$F^* = f(\mathbf{X}^*(\mathbf{p}), \mathbf{p}) \quad (14)$$

En ocasiones la obtención de la solución óptima lleva a meditar sobre la conveniencia de modificar alguno de los valores de \mathbf{p} fijados previamente y a considerar cual hubiera sido la solución para otros valores distintos. Dado que habitualmente las variables de diseño son funciones implícitas de \mathbf{p} no es posible obtener la nueva solución directamente de (13) y si se modifica algún parámetro p_k el proceso de optimización debe reiniciarse. Sin embargo otra alternativa sería linealizar las expresiones (13) y (14) y escribir:

$$F = F(\mathbf{X}^*) + \sum \frac{dF(\mathbf{X}^*)}{dp_k} \Delta p_k \quad (15)$$

$$\mathbf{X} = \mathbf{X}^* + \frac{d\mathbf{X}^*}{dp_k} \Delta p_k \quad (16)$$

Las expresiones $dF(\mathbf{X}^*)/dp_k$ y $d\mathbf{X}^*/dp_k$ son las sensibilidades de primer orden de la función objetivo y las variables de diseño respecto al parámetro p_k .

Los métodos propuestos para obtener estas sensibilidades son varios y pueden destacarse los siguientes, propuestos por Vanderplaats⁵, Sobieszcanski-Sobieski⁶ y Barthelemy⁷:

- Métodos basados en aplicación de las condiciones de Kuhn-Tucker.
- Métodos en utilización de la función penalty.
- Métodos basados en las direcciones eficientes.

En las referencias indicadas se desarrollan cada uno de esos métodos. A modo de ejemplo se desarrollará cómo funciona el método basado en las condiciones de Kuhn-Tucker y se aplicará el conocido ejemplo⁸ de celosía de tres barras de la Figura 3.

Las condiciones de Kuhn-Tucker expresan que en el óptimo existe un conjunto de condiciones activas con las que se cumple.

$$\nabla F(\mathbf{X}^*) + \nabla G(\mathbf{X}^*)\lambda = \mathbf{0} \quad (17.a)$$

$$g_j(\mathbf{X}^*) = 0 \quad j = 1, \dots, J \quad (17.b)$$

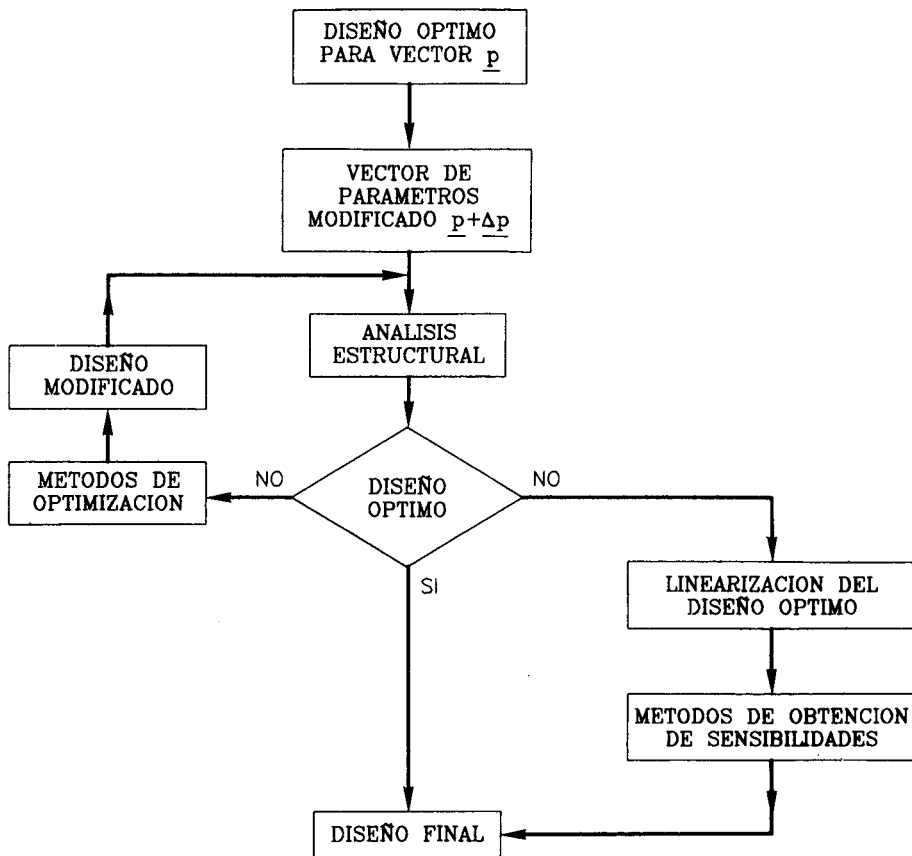


Figura 9. Diagrama de flujo de las alternativas en diseño óptimo de estructuras, en caso de modificaciones en el vector \underline{p} de parámetros.

$$\lambda_j \geq 0 \quad (17.c)$$

siendo λ el vector de multiplicadores de Lagrange y ∇G una matriz cuyas columnas son los gradientes de las condiciones $g_j(\mathbf{X})$.

$$\nabla G = \left| \nabla g_1(\mathbf{X}^*), \dots, \nabla g_J(\mathbf{X}^*) \right| \quad (18)$$

Derivando (17.a) y (17.b) respecto a un parámetro p_k y escribiendo matricialmente la expresión se tiene:

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{vmatrix} \begin{vmatrix} d\mathbf{X} \\ d\lambda \end{vmatrix} + \begin{vmatrix} \mathbf{C} \\ \mathbf{D} \end{vmatrix} = \mathbf{0} \quad (19)$$

los elementos de cada matriz son:

$$a_{il} = \frac{\partial^2 F(\mathbf{X}^*)}{\partial x_i \partial x_l} + \sum_{j=1}^J \lambda_j \frac{\partial^2 g_j(\mathbf{X}^*)}{\partial x_i \partial x_l} \quad (20.a)$$

$$b_{ij} = \frac{\partial g_j(\mathbf{X}^*)}{\partial x_i} \quad (20.b)$$

$$c_i = \frac{\partial^2 F(\mathbf{X}^*)}{\partial x_i \partial p_k} + \sum_{j=1}^J \lambda_j \frac{\partial^2 g_j(\mathbf{X}^*)}{\partial x_i \partial p_k} \quad (20.c)$$

$$d_j = \frac{\partial g_j(\mathbf{X}^*)}{\partial p_k} \quad (20.d)$$

El vector λ puede obtenerse de la expresión:

$$\lambda = - \left| \nabla \mathbf{G}^T \nabla \mathbf{G} \right|^{-1} \nabla \mathbf{G}^T \nabla \mathbf{F} \quad (21)$$

La solución del sistema (19) proporciona las sensibilidades $d\lambda$ y $d\mathbf{X}$. Por su parte la sensibilidad de la función objetivo será:

$$\frac{dF}{dp_k} = \frac{\partial F}{\partial p_k} + \sum_{i=1}^n \frac{\partial F}{\partial x_i} \frac{dx_i}{dp_k} \quad (22)$$

Este método presupone que aunque se modifique el valor de algún parámetro se mantiene el subgrupo de condiciones activas, pero eso será cierto únicamente en un entorno del valor inicial del parámetro p_k . Como al hacerse pasiva la condición el multiplicador de Lagrange asociado se anula, ello nos da la pauta para obtener el incremento necesario del parámetro, que resultará de:

$$\lambda_j(p_k) + \frac{d\lambda_j}{dp_k} \Delta p_k = 0 \quad (23)$$

Si el método anterior se aplica a la celosía de tres barras de la Figura 3 las expresiones (17) resultan:

$$\frac{\partial F}{\partial x_1} + \lambda \frac{\partial \sigma_1}{\partial x_1} = 0 \quad (24.a)$$

$$\frac{\partial F}{\partial x_2} + \lambda \frac{\partial \sigma_1}{\partial x_2} = 0 \quad (24.b)$$

$$g_1 = \sigma_1 - \sigma_M = 0 \quad (24.c)$$

siendo

$$F = (2\sqrt{2}x_1 + x_2)l \quad (25.a)$$

$$g_1 = \frac{x_2 + \sqrt{2}x_1}{2x_1x_2 + \sqrt{2}x_1^2} P - \sigma_M = 0 \quad (25.b)$$

Los elementos de las matrices (20) son:

$$\frac{\partial^2 F}{\partial x_i \partial x_l} = \frac{\partial^2 F}{\partial x_i \partial \sigma_M} = \frac{\partial^2 g_1}{\partial x_i \partial \sigma_M} = 0 \quad (26.a)$$

$$\frac{\partial g_1}{\partial x_1} = -\frac{x_1^2 + \sqrt{2}x_1x_2 + x_2^2}{(x_1^2 + \sqrt{2}x_1x_2)^2}P \quad \frac{\partial g_1}{\partial x_2} = -\frac{\sqrt{2}x_1^2}{(\sqrt{2}x_1^2 + 2x_1x_2)^2}P \quad (26.b)$$

$$\frac{\partial^2 g_1}{\partial x_1^2} = \frac{2x_1^3 + 3\sqrt{2}x_1^2x_2 + 6x_1x_2^2 + 2\sqrt{2}x_2^3}{(x_1^2 + \sqrt{2}x_1x_2)^3}P \quad \frac{\partial^2 g_1}{\partial x_2^2} = -\frac{2P}{(x_1 + \sqrt{2}x_2)^3} \quad (26.c)$$

$$\frac{\partial^2 g_1}{\partial x_1 \partial x_2} = \frac{\sqrt{2}P}{(x_1 + \sqrt{2}x_2)^3} \quad (26.d)$$

$$\frac{\partial g_1}{\partial \sigma_M} = -1 \quad (26.e)$$

La solución resultante es

$$\frac{dx_1}{d\sigma_M} = -0.788 \frac{P}{\sigma_M^2} \quad \frac{dx_2}{d\sigma_M} = -0.408 \frac{P}{\sigma_M^2} \quad \frac{d\lambda}{d\sigma_M} = -5.278 \frac{Pl}{\sigma_M^3} \quad (27)$$

y la sensibilidad de la función objetivo

$$\frac{dF}{d\sigma_M} = \left(2\sqrt{2} \frac{dx_1}{d\sigma_M} + \frac{dx_2}{d\sigma_M} \right) l = -2.639 \frac{Pl}{\sigma_M^2} \quad (28)$$

En consecuencia la solución óptima asociada a otro valor de la tensión admisible se puede aproximar a:

$$F = F^* + \frac{dF^*}{d\sigma_M} \Delta\sigma_M \quad (29)$$

$$x_i = x_i^* + \frac{dx_i^*}{d\sigma_M} \Delta\sigma_M \quad i = 1, 2 \quad (30)$$

Supóngase que

$$P = 2 \text{ t} \quad \sigma_M = 2 \text{ t/cm}^2 \quad l = 100 \text{ cm}$$

La tabla adjunta presenta los resultados correspondientes a la solución exacta y a la obtenida por esta aproximación lineal para los incrementos $\Delta\sigma_M = \pm 10\%$ y $\Delta\sigma_M = \pm 20\%$.

$\Delta\sigma_M$	x_1	x_2	F	x_1	x_2	F	error
-20%	0.98	0.51	329.9	0.94	0.49	316.6	4%
-10%	0.87	0.45	293.2	0.86	0.45	290.3	1%
10%	0.72	0.37	239.9	0.71	0.37	237.5	1%
20%	0.66	0.34	219.9	0.63	0.32	211.1	4%

Tabla I.

REFERENCIAS

1. A. Osyczka, "Multicriterion Optimization in Engineering with FORTRAN Programs", Ellis Herwood, (1984).
2. H. Eschenauer, J. Koski y A. Osyczka (ed.), "Multicriteria Design Optimizations. Procedures and Applications", Springer Verlag, (1990).
3. V. Pareto, "Cours D'Economie Politique", Vol I y II, F. Rouge, (1886).
4. L. Duckstein, "Multiobjective Optimization in Structural Design: The Model Choice Problem", *New Directions in Optimum Design*, Atrek, E. y otros (ed.) John Wiley, (1984).
5. G.N. Vanderplaats, y H. Yoshida, "Efficient Calculation of Optimum Design Sensitivity", *AIAA Journal*, Vol. 23, pp. 1798-1803, (1985).
6. J. Sobieszczansky-Sobieski, J.F.M. Barthelemy y K.M. Riley, "Sensitivity of Optimum Solutions of Problem Parameters", *AIAA Journal*, Vol. 20, pp. 1291-1299, (1982).
7. J.F.M. Barthelemy y J. Sobieszczansky-Sobieski, "Sensitivity Derivatives of Objective Functions in Nonlinear Programming", *AIAA Journal*, Vol. 21, pp. 913-915, (1983).
8. S. Hernández y J. Mata, "Multicriteria Optimization and Global Optimization for Training in Structural Design", *CATS'90, Congreso Internacional en Computer Aided Training in Science and Technology*, E. Oñate y otros (ed.), CIMNE-Pineridge Press, pp. 685-628, (1990).